

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2014-137635
(P2014-137635A)

(43) 公開日 平成26年7月28日(2014.7.28)

(51) Int. Cl. F I テーマコード (参考)
G06F 11/30 (2006.01) G06F 11/30 310A 5B042
G06F 11/34 (2006.01) G06F 11/34 S

審査請求 未請求 請求項の数 3 O L (全 19 頁)

(21) 出願番号 特願2013-4728 (P2013-4728)
 (22) 出願日 平成25年1月15日 (2013.1.15)

(71) 出願人 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番1号
 (74) 代理人 100119161
 弁理士 重久 啓子
 (74) 代理人 100111822
 弁理士 渡部 章彦
 (72) 発明者 金政 泰彦
 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
 (72) 発明者 久保田 敦
 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

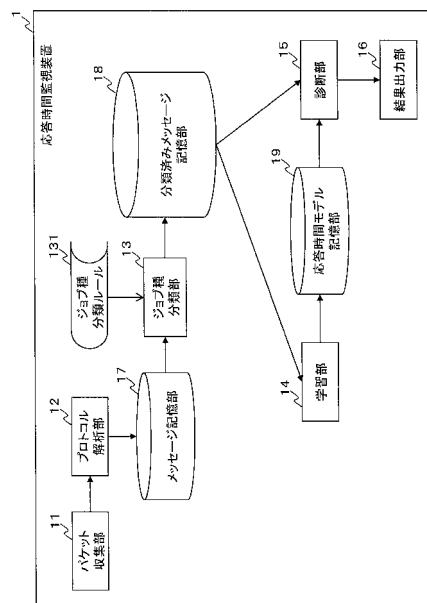
(54) 【発明の名称】 応答時間監視プログラム、方法および応答時間監視装置

(57) 【要約】

【課題】 応答時間監視装置で、監視対象が実行するジョブに最適な閾値を用いて応答時間を監視することを目的とする。

【解決手段】 応答時間監視装置 1 は、学習部 14 が、監視対象コンピュータで一定時間に実行される全ジョブ合計の同時実行数と各ジョブ種でのジョブ種単独の同時実行数と平均応答時間とを計算し、ジョブ種毎に、その平均応答時間と全ジョブ合計の同時実行数との第 1 の対応関係とジョブ種単独の同時実行数との第 2 の対応関係を生成し、該ジョブ種に対応する第 1 の対応関係および第 2 の対応関係のうち平均応答時間との相関が高い対応関係を選択する学習部 14 と、予め設定された時間毎に前記監視対象コンピュータで実行される各ジョブの応答時間を取得し、各ジョブの応答時間を、該ジョブのジョブ種に対応して選択された対応関係を用いて決定した閾値をもとにアラート対象とするかを判断する診断部 15 を備える。

【選択図】 図 4



【特許請求の範囲】**【請求項 1】**

監視対象コンピュータの応答時間を監視するために、コンピュータに、
監視対象コンピュータで一定時間に実行される全ジョブ合計の同時実行数と、各ジョブ種について、ジョブ種単独の同時実行数と該ジョブ種の平均応答時間とを計算し、
前記ジョブ種毎に、該ジョブ種の平均応答時間と全ジョブ合計の同時実行数との第 1 の対応関係と、該ジョブ種の平均応答時間とジョブ種単独の同時実行数との第 2 の対応関係とを生成し、
前記ジョブ種毎に、対応する前記第 1 の対応関係および前記第 2 の対応関係のうち、該ジョブ種の平均応答時間との相関が高い対応関係を選択し、
予め設定された時間毎に前記監視対象コンピュータで実行される各ジョブの応答時間を取得し、
前記各ジョブの応答時間を、該ジョブのジョブ種に対応して選択された対応関係を用いて決定した閾値をもとにアラート対象とするかを判断する、処理を実行させるための
応答時間監視プログラム。

10

【請求項 2】

監視対象コンピュータの応答時間を監視するために、コンピュータが、
監視対象コンピュータで一定時間に実行される全ジョブ合計の同時実行数と、各ジョブ種について、ジョブ種単独の同時実行数と該ジョブ種の平均応答時間とを計算し、
前記ジョブ種毎に、該ジョブ種の平均応答時間と全ジョブ合計の同時実行数との第 1 の対応関係と、該ジョブ種の平均応答時間とジョブ種単独の同時実行数との第 2 の対応関係とを生成し、
前記ジョブ種毎に、対応する前記第 1 の対応関係および前記第 2 の対応関係のうち、該ジョブ種の平均応答時間との相関が高い対応関係を選択し、
予め設定された時間毎に前記監視対象コンピュータで実行される各ジョブの応答時間を取得し、
前記各ジョブの応答時間を、該ジョブのジョブ種に対応して選択された対応関係を用いて決定した閾値をもとにアラート対象とするかを判断する、処理を実行する
応答時間監視方法。

20

【請求項 3】

監視対象コンピュータで一定時間に実行される全ジョブ合計の同時実行数と、各ジョブ種について、ジョブ種単独の同時実行数と該ジョブ種の平均応答時間とを計算し、前記ジョブ種毎に、該ジョブ種の平均応答時間と全ジョブ合計の同時実行数との第 1 の対応関係と、該ジョブ種の平均応答時間とジョブ種単独の同時実行数との第 2 の対応関係とを生成し、前記ジョブ種毎に、対応する前記第 1 の対応関係および前記第 2 の対応関係のうち、該ジョブ種の平均応答時間との相関が高い対応関係を選択する学習部と、
予め設定された時間毎に前記監視対象コンピュータで実行される各ジョブの応答時間を取得し、前記各ジョブの応答時間を、該ジョブのジョブ種に対応して選択された対応関係を用いて決定した閾値をもとにアラート対象とするかを判断する診断部とを備える
応答時間監視装置。

30

40

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、コンピュータの応答時間監視処理に関する。

【背景技術】**【0002】**

監視対象であるコンピュータの応答時間を監視し、正常と許容できる範囲（正常範囲）を超えた応答時間となった場合にアラートを発するが、正常範囲の閾値（上限値、下限値）を適切に設定することは難しい。

50

【0003】

正常範囲の閾値が低過ぎる場合は、アラートが多発して管理者がアラートに慣れてしまい注意を払わなくなるおそれがある。また、閾値が高過ぎる場合は、問題が生じていてもアラートが発せられない状態となり、管理者が問題に対処するのが遅れてしまうおそれがある。

【0004】

さらに、コンピュータの短期的な負荷集中によって自然に応答時間が増加したような場合は、管理者が対処を要する問題が発生していないので、アラートを抑制することが望ましい。

【0005】

例えば、電子メール管理サーバ(コンピュータ)の応答時間が特定の時間帯に増加する場合、人事管理システムや給与管理システムのコンピュータの応答時間が特定の期間に増加する場合など、一時的な処理集中による応答時間の増加に基づくアラートが発せられると、管理者はアラートの原因の調査と、対応策が要求されるため、管理業務の負担になるだけでなく、不要なアラートに対する慣れによって本来のアラートを見落とす可能性が高くなる等の問題が生じる。

【0006】

そのため、従来では、管理者は、応答時間の増加によるアラートがあると、監視対象のコンピュータに対するアラートの原因調査と発生している問題への対処が必要であるかの判断を手作業で行っていた。詳しくは、監視対象のコンピュータの応答時間についてアラートが検出されると、管理者が他の様々な性能指標をチェックし、コンピュータに何が起きているのかを総合的に判断し、問題が生じていると判断した場合に必要な対処を行っていた。

【0007】

また、従来、複数のランクの警報を発するために、監視対象の装置データ間の関係を示す不変量を抽出して、複数のルールの複数の等価しきい値を不変量のネットワークで計算し、監視対象の装置から警報を受信すると、普段の状態とどの程度異なる状況かによりその警報をランク付ける装置が知られている。

【0008】

さらに、従来、分散システムの能力計画と資源最適化のために、各装置や資源から収集された測定値間の関係を示す不変式を求め、不変式とユーザ負荷を示すフロー強度とにより分散システムを構成する要素の能力を決定する方法が知られている。

【先行技術文献】

【特許文献】

【0009】

【特許文献1】特表2011-521380号公報

【特許文献2】特表2010-507146号公報

【発明の概要】

【発明が解決しようとする課題】

【0010】

監視対象のコンピュータに発生した問題による応答時間の増加の場合は、速やかにアラートを発生させる必要があるが、上記のように、コンピュータの短期的な負荷集中によって自然に応答時間が増加した場合にはアラートを抑制することが望ましい。

【0011】

コンピュータの負荷の集中による応答時間の増加する場合として、一般的に、多数のユーザが同時にアクセスする場合があるが、時間経過により自然解消するためアラートが不要である。また、排他ロックを行う処理が実行されている場合には、アクセスするユーザ数が少なくても応答時間が一時的に増加するが、アラートは不要である。

【0012】

そのため、短期的な負荷集中による自然な応答時間の増加と問題発生による異常な応答

10

20

30

40

50

時間の増加とを区別してアラートを発生させられるような正常範囲の設定が必要であるが、従来手法では、2つの応答時間の増加を区別できる正常範囲の設定が難しい。

【0013】

例えば、複数データ間の相関分析にもとづいて異常を判断する場合には、正常範囲の設定に用いる複数データの相関が不明であり、自然な負荷集中であっても相関が乱れただけで異常と判断されてアラートが多発する可能性がある。

【0014】

さらに、監視対象に実際に問題が発生することなく、応答時間の正常範囲の閾値を適切に設定する必要がある。

【0015】

本発明の目的は、監視対象のコンピュータの応答時間について、一時的な負荷集中と問題発生とを区別できるように正常範囲を設定し、異常な応答時間の増加のみにアラートを出力できる応答時間監視技術を提供することである。

【課題を解決するための手段】

【0016】

本発明の一態様として開示され応答時間監視プログラムは、監視対象コンピュータの応答時間を監視するために、コンピュータに、1)監視対象コンピュータで一定時間に実行される全ジョブ合計の同時実行数と、各ジョブ種について、ジョブ種単体の同時実行数と該ジョブ種の平均応答時間とを計算し、2)前記ジョブ種毎に、該ジョブ種の平均応答時間と全ジョブ合計の同時実行数との第1の対応関係と、該ジョブ種の平均応答時間とジョブ種単体の同時実行数との第2の対応関係を生成し、3)前記ジョブ種毎に、対応する前記第1の対応関係および前記第2の対応関係のうち、該ジョブ種の平均応答時間との相関が高い対応関係を選択し、4)予め設定された時間毎に前記監視対象コンピュータで実行される各ジョブの応答時間を取得し、5)前記各ジョブの応答時間を、該ジョブのジョブ種に対応して選択された対応関係を用いて決定した閾値をもとにアラート対象とするかを判断する、処理を実行させるためのものである。

【発明の効果】

【0017】

開示する応答時間監視プログラムによれば、監視対象の応答時間の増加が短期的な負荷集中による自然な応答時間増加であるか、問題の発生による異常な応答時間増加であるかを区別した監視を実現することができる。

【図面の簡単な説明】

【0018】

【図1】図1は、応答時間監視装置の一実施例におけるハードウェア構成例を示す図である。

【図2】同時実行数と応答時間の関係の直線近似をもとに応答時間の正常範囲を設定する根拠を説明するための図である。

【図3】応答時間監視装置が実施されるシステムの構成例を示す図である。

【図4】応答時間監視装置の一実施例における機能ブロック構成例を示す図である。

【図5】一実施例においてメッセージ記憶部に記憶されているプロトコルメッセージの例を示す図である。

【図6】一実施例におけるHTTPプロトコルのジョブ種の分類ルールを示す図である。

【図7】一実施例におけるDBプロトコルのジョブ種の分類ルールを示す図である。

【図8】一実施例における一定の時間間隔に分割した各細分化区間および各ジョブの応答時間の例を示す図である。

【図9】一実施例における同時実行数と応答時間の計算例を示す図である。

【図10】一実施例における同時実行数と応答時間の対応関係のプロット例を示す図である。

【図11】一実施例における直線近似および同時実行数の選択を説明するための図である。

。

10

20

30

40

50

【図 1 2】一実施例におけるグラフ上での信頼区間による応答時間の正常範囲の設定例を示す図である。

【図 1 3】一実施例における応答時間モデルの例を示す図である。

【図 1 4】一実施例における学習処理の処理フロー例を示す図(1)である。

【図 1 5】一実施例における学習処理の処理フロー例を示す図(2)である。

【図 1 6】一実施例における応答時間の診断処理の処理フロー例を示す図である。

【図 1 7】一実施例におけるステップ S 3 2 の診断処理のより詳細な処理フロー例を示す図である。

【発明を実施するための形態】

【0019】

以下、本発明の一態様として開示する応答時間監視方法を実行する応答時間監視装置について説明する。

【0020】

図 1 は、応答時間監視装置 1 の一実施例におけるハードウェア構成例を示す図である。

【0021】

応答時間監視装置 1 は、CPU 101、短期記憶部(DRAM)102、長期記憶部(HDD)103、ネットワークインタフェース104、入力装置(キーボード、マウス等)105、出力装置(ディスプレイ、プリンタ等)106が内部ネットワーク等で接続されたコンピュータとして実施することができる。

【0022】

応答時間監視装置 1 は、監視対象のコンピュータ、コンピュータシステム等の応答時間監視処理に必要な情報をファイルとして長期記憶部 103 に記憶し、入力装置 105 から実行プログラムを起動し、起動された実行プログラムが、短期記憶部 102 にロードされ、ネットワークインタフェース 104 で受信した監視対象がパケットデータをもとに処理を実行する。

【0023】

応答時間監視装置 1 は、必要に応じて情報を長期記憶部 103 から短期記憶部 102 に読み出しながら応答時間監視処理を進め、監視対象の応答時間が正常範囲を超えた場合に、ネットワークインタフェース 104 から予め設定された管理者の端末装置へアラート情報を送信する。また、応答時間監視装置 1 は、アラート情報を出力装置 106 へ出力することもできる。

【0024】

なお、応答時間監視処理の実行プログラムは、CD-ROM、CD-RW、DVD-R、DVD-RAM、DVD-RW等やフレキシブルディスク等の記録媒体だけでなく、通信回線の先に備えられた他の記憶装置やコンピュータのハードディスク等に記憶されるものであってもよい。

【0025】

応答時間監視装置 1 は、監視対象のコンピュータまたはコンピュータシステムの応答時間が正常であると許容できる範囲(正常範囲)を示す閾値(上限値または下限値またはその両方)を監視対象のコンピュータの負荷量に基づいて動的に変更する。

【0026】

応答時間監視装置 1 が監視する応答時間は、監視対象のコンピュータまたはコンピュータシステムにリクエストが到達した時点からレスポンスが出力する時点までの経過時間を意味する。

【0027】

また、監視対象の負荷量として、一定の時間単位毎のジョブの同時実行数を用いる。

【0028】

コンピュータで同時期に実行されているジョブ数(同時実行数)の増加に応じて応答時間が増加し、ハードウェア資源が飽和すると応答時間はジョブの同時実行数の増加分を超えて増加する。

10

20

30

40

50

【 0 0 2 9 】

応答時間監視装置 1 は、監視対象のコンピュータ上で実行されるジョブ種毎の応答時間と共に、一定時間内に実行されているジョブ種毎（ジョブ種単独）の同時実行数と全てのジョブ（全ジョブ合計）の同時実行数とを計測する。

【 0 0 3 0 】

そして、応答時間監視装置 1 は、学習フェーズとして、ジョブ種毎に、その応答時間と 2 つの種類の同時実行数それぞれとの対応関係を求め、全ジョブ合計とジョブ種単独とどちらの種類の同時実行数とより相関するかを評価し、より相関すると評価した同時実行数の種類と応答時間との対応関係の直線近似を求め、近似直線から統計上の信頼区間の上限および下限を応答時間の正常範囲として、その閾値（上限値、下限値）を示す応答時間モデルを求めて保持しておく。

10

【 0 0 3 1 】

図 2 は、同時実行数と応答時間の関係の直線近似をもとに応答時間の正常範囲を設定する根拠を説明するための図である。

【 0 0 3 2 】

図 2 において（ a ）で示す同時実行数の分布は、個々のジョブが消費する資源（CPU 等）の量のばらつきを示している。例えば、同じジョブ種であっても具体的なジョブ内容によって、処理に必要な資源量は異なる。また、同時実行数の増加に伴う応答時間の増加も理論値からのばらつきが生じる。図 2 において（ b ）で示す傾きは、競合している資源が単一であって他の資源の待ち時間が発生しない範囲では、ジョブの応答時間の理論値が同時実行数に比例して増加する傾向を示している。

20

【 0 0 3 3 】

図 2 において（ c ）で示す近似直線からの増分は、資源利用間の依存関係によって発生する他の競合資源の待ち時間を示しており、ジョブの同時実行数が多くなると発生しやすくなる。

【 0 0 3 4 】

図 2 において（ d ）で示す部分は、ある資源が完全に飽和した場合には、それ以上の同時実行が処理しきれないことにより、ジョブの同時実行数の増加が待ち時間の増加にそのまま繋がり、生じた待ち時間によって応答時間が等比級数的に増加することを示している。したがって、（ d ）の部分のように、計測された応答時間が近似直線から信頼区間を超える場合は、資源が完全に飽和しているとみなすことができ、監視対象での問題発生による応答時間の増加として管理者へのアラートが必要となる。応答時間監視装置 1 は、図 2 の（ a ）～（ c ）に該当する応答時間と（ d ）の部分に該当する応答時間とを区別して、（ d ）に該当する応答時間の増加をアラート対象とする。

30

【 0 0 3 5 】

その後、応答時間監視装置 1 は、診断フェーズとして、一定時間毎に全ジョブの開始時刻および終了時刻を取得して、ジョブ種毎に、平均応答時間を計算し、学習結果である応答時間モデルで定められた種別の同時実行数と平均応答時間から正常範囲の閾値を決定し、計算した応答時間が閾値を超える場合にそのジョブ種をアラート対象（応答時間が正常範囲を超えたジョブ種）として出力する。

40

【 0 0 3 6 】

図 3 は、応答時間監視装置 1 が実施されるシステムの構成例を示す図である。

【 0 0 3 7 】

応答時間監視装置 1 の監視対象は、ウェブ（Web）サーバ 2 1、データベース（DB）サーバ 2 2 等を含むコンピュータシステム（以下、監視対象システムと呼ぶ）2 である。監視対象システム 2 の Web サーバ 2 1、DB サーバ 2 2 はネットワークスイッチ 3 に接続する。Web サーバ 2 1 は、さらにプロキシ（ファイアウォール）4 を介してインターネット 5 に接続し、クライアント 6 のリクエストに対してレスポンスを返してサービスを提供する。

【 0 0 3 8 】

50

応答時間監視装置 1 は、監視対象システム 2 とプロキシ 4 との間に設けられたネットワークスイッチ 3 に接続する。また、応答時間監視装置 1 は、ネットワークスイッチ 3 に接続する性能監視サーバ 7 の内部機能として実施されてもよい。

【0039】

応答時間監視装置 1 は、ネットワークスイッチ 3 が持つポートミラーリング機能を用いて、監視対象システム 2 の Webサーバ 2 1 - DBサーバ 2 2 間、Webサーバ 2 1 - クライアント 6 間で送受信されるパケットデータ (IP パケット) を収集する。

【0040】

図 4 は、応答時間監視装置 1 の一実施例における機能ブロック構成例を示す図である。

【0041】

応答時間監視装置 1 は、応答時間監視処理を実行するために、パケット収集部 1 1、プロトコル解析部 1 2、ジョブ種分類部 1 3、学習部 1 4、診断部 1 5、結果出力部 1 6 を備え、さらに、データ保管場所として、メッセージ記憶部 1 7、分類済みメッセージ記憶部 1 8、応答時間モデル記憶部 1 9 を備える。

【0042】

パケット収集部 1 1 は、監視対象システム 2 の Webサーバ 2 1 - DBサーバ 2 2 間、Webサーバ 2 1 - クライアント 6 間で所定の時間内に送受信されたパケットデータ (IP パケット) のストリームをネットワークスイッチ 3 から収集し、各パケットデータに取得時刻を付加する。応答時間監視装置 1 が性能監視サーバ 7 の内部機能として実施される場合は、パケット収集部 1 1 は、性能監視サーバ 7 のタイムスタンプ機能によりタイムスタンプ (取得時刻) が付加されたパケットデータを収集する。

【0043】

プロトコル解析部 1 2 は、パケット収集部 1 1 が収集したパケットデータのプロトコルを解析し、メッセージの種類 (リクエスト (request)、レスポンス (response) 等)、パケットデータの送信元および送信先の IP アドレス、ポート番号、セッション番号、コマンド名、CGI のパラメータ等を抽出し、抽出した情報と取得時刻とを含むプロトコルメッセージをメッセージ記憶部 1 7 に記憶する。

【0044】

図 5 は、メッセージ記憶部 1 7 に記憶されているプロトコルメッセージの例を示す図である。

【0045】

各プロトコルメッセージは、パケットの取得時刻、セッション番号、送信元および送信先の IP アドレスおよびポート番号、コマンド名、メッセージの種類等を含む。例えば、図 5 に示す番号 = 1 のプロトコルメッセージは、パケットの取得時刻「2009/09/07 12:12:04.787360」、セッション番号「132290-1」、メッセージの送信元の IP アドレス「194.185.39.24」とポート番号「51272」、送信先の IP アドレス「194.23.5.226」とポート番号「10443」、プロトコルの種類「Request HTTP」を含む。

【0046】

ジョブ種分類部 1 3 は、プロトコルメッセージをジョブ毎に分類してリクエストとレスポンスのメッセージを紐付け、ジョブ種分類ルール 1 3 1 に基づいてジョブ種に分類したプロトコルメッセージを分類済みメッセージ記憶部 1 8 に記憶する。

【0047】

図 5 に示すプロトコルメッセージの例では、番号 = 1、20 のリクエストとレスポンスのプロトコルメッセージが同一ジョブとして紐付けられ、番号 9、10 のリクエストとレスポンスのプロトコルメッセージが同一ジョブとして紐付けられる。

【0048】

ジョブ種分類ルール 1 3 1 は、プロトコルメッセージに含まれるジョブ内容からジョブ種を特定するための規則を示す情報である。

【0049】

Webサーバ 2 1 とクライアント 6 の間では HTTP メッセージが送受信されている。

10

20

30

40

50

HTTPメッセージは、「http://www.server.com/job/type.jsp?param1=foo¶m2=bar」のような識別子でリクエストされ、Webサーバ21（www.server.comというサーバ）ではリクエストに対して「/job/type.jsp」のようなアドレス表記で指定されるコンテンツをレスポンスとして返却する。

【0050】

Webサーバ21とDBサーバ22の間ではSQLによる問い合わせとその問い合わせ結果が送受信されている。例えば「SELECT 従業員番号 FROM 従業員テーブル WHERE 従業員ID=00001」という問い合わせに対して、DBサーバ22は、「従業員テーブル」から「従業員ID=00001」という条件にあうデータを検索し、問い合わせ結果をWebサーバ21に返却する。

10

【0051】

ジョブ種分類部13は、HTTPプロトコルのジョブ種では、ジョブ種分類ルール131に基づいて、アドレス毎に1種類、またはアドレスと一部のCGIパラメータとの組み合わせ（例えば、/job/type.jsp?param2=ba）からジョブ種を決める。

【0052】

図6は、HTTPプロトコルのジョブ種の分類ルールを示す図である。

【0053】

HTTPプロトコルのジョブ種分類ルール131は、コマンド名、アドレス内のローカルアドレスの一部、CGIパラメータの一部を用いて各ジョブ種を特定することを表している。

20

【0054】

また、ジョブ種分類部13は、DBプロトコルのジョブ種では、ジョブ種分類ルール131に基づいてSQLを抽象化したもの、例えば「従業員テーブルにアクセスするSELECT文」のようにしたものを1つのジョブ種として決める。

【0055】

図7は、DBプロトコルのジョブ種の分類ルールを示す図である。

【0056】

DBプロトコルメッセージのジョブ種分類ルール131は、コマンド名、およびSQL文の内容を抽象化する変換ルールでジョブ種を特定することを表している。

【0057】

30

学習部14は、分類済みメッセージ記憶部18に記憶されたジョブ種で分類されたプロトコルメッセージを学習データとして入力し、ジョブ種毎に応答時間の正常範囲を示す閾値を示す応答時間モデル（モデル式）を求める。

【0058】

診断部15は、予め設定された一定時間毎に、分類済みメッセージ記憶部18に記憶されたジョブ種で分類されたプロトコルメッセージのデータストリームを入力し、応答時間モデル記憶部19に記憶された応答時間モデルを用いて、計測された各ジョブ種の応答時間が正常範囲内であるかを判断する。

【0059】

結果出力部16は、診断部15が、正常範囲を超えた応答時間と判断したジョブ種をアラートとして出力する。

40

【0060】

メッセージ記憶部17は、収集した取得時刻付のプロトコルメッセージを記憶する。

【0061】

分類済みメッセージ記憶部18は、ジョブ種で分類された取得時刻付のプロトコルメッセージを記憶する。

【0062】

応答時間モデル記憶部19は、学習部14で得た各ジョブ種の応答時間の正常範囲を示す応答時間モデルを記憶する。

【0063】

50

次に、応答時間監視装置 1 の学習部 1 4 の学習処理（学習フェーズ）をより詳しく説明する。

【0064】

学習部 1 4 は、分類済みメッセージ記憶部 1 8 に記憶された一定の時間に収集されたプロトコルメッセージを学習データとし、以下の処理を行う。

【0065】

(1) 学習データを予め定めた一定の時間間隔に分割する。さらに、ジョブ毎に紐付けたリクエストのプロトコルメッセージの取得時刻を開始時刻、レスポンスのプロトコルメッセージの取得時刻を終了時刻として各ジョブの応答時間を計算する。

【0066】

図 8 は、一定の時間間隔に分割した各細分化区間および各ジョブの応答時間の例を示す図である。

【0067】

図 8 において、矩形は 1 つのジョブを示し、矩形の下の値は当該ジョブの応答時間（ミリ秒 [ms]）、矩形内の D_n (n は 1、2、...) は当該ジョブのジョブ種を表している。

【0068】

(2) 学習データを予め定めた時間間隔に分割し、分割した区間（細分化区間：100ms）内で、全ジョブ合計の平均同時実行数、ジョブ種単独の平均同時実行数、各ジョブ種の平均応答時間を、以下の式で計算する。

【0069】

全ジョブ合計の平均同時実行数 = 全ジョブの応答時間の合計 / 細分化区間
 ジョブ種単独の平均同時実行数 = 当該ジョブ種の応答時間 / 細分化区間
 各ジョブ種の平均応答時間 = 当該ジョブ種の全応答時間 / 当該ジョブ種の平均同時実行数

なお、2 つの細分化区間にまたがるジョブは、応答時間を各細分化区間で案分してそれぞれの応答時間とする。

【0070】

図 9 は、同時実行数と応答時間の計算例を示す図である。

【0071】

学習データの各ジョブの応答時間が図 7 に示す状態である場合に、それぞれ以下の計算となる。例えば、

$$\begin{aligned} \text{全ジョブ合計の平均同時実行数} \\ &= (12 + 11 + 58 + 11 + 15 + 30 + 38 + 9 + 7 + 10) / 100 \\ &= 2.01、 \end{aligned}$$

$$\begin{aligned} \text{ジョブ種 D 1 のジョブ種単独の平均同時実行数} \\ &= (11 + 15 + 9) / 100 = 0.35、 \end{aligned}$$

$$\begin{aligned} \text{ジョブ種 D 1 の各ジョブ種の応答時間} \\ &= (11 + 15 + 9) / 3 = 11.7 \end{aligned}$$

となる。

【0072】

ジョブ種 D 2 の場合は、3 番目に実行されたジョブが 2 つの細分化区間 T 1、T 2 で案分されるため、

$$\begin{aligned} \text{ジョブ種 D 2 のジョブ種単独の平均同時実行数} \\ &= (30 + 38 + 10) / 100 = 0.78、 \\ \text{ジョブ種 D 2 の各ジョブ種の平均応答時間} \\ &= (30 + 38 + 10) / 1 + 1 + 10 / 34 = 34.0 \end{aligned}$$

となる。

【0073】

(3) 上記(2)の計算結果から、細分化区間数分の「全ジョブ合計の平均同時実行数

10

20

30

40

50

、当該ジョブ種の応答時間」の第1の組データと、「当該ジョブ種のジョブ種単独の同時実行数、当該ジョブ種の平均応答時間」の第2の組データとを生成する。

【0074】

(4) 各ジョブ種について、生成した2種の組データそれぞれに対応するグラフを用意し、各細分化区間での平均同時実行数と当該ジョブ種の平均応答時間との対応関係を各グラフ上にプロットする。

【0075】

図10は、同時実行数と応答時間の対応関係のプロット例を示す図である。

【0076】

図10の左側に示すグラフは、全ジョブ合計の同時実行数を横軸に応答時間を縦軸とするグラフであり、ジョブ種D1の第1の組データに基づく各細分区間での平均同時実行数と平均応答時間との対応関係をプロットしたものである。

10

【0077】

図10の右側に示すグラフは、ジョブ種単独の同時実行数を横軸に応答時間を縦軸にとるグラフであり、ジョブ種D1の第2の組データに基づく各細分化区間での平均同時実行数と平均応答時間との対応関係をプロットしたものである。

【0078】

図10に示す2種のグラフが各ジョブ種について生成される。

【0079】

(5) 上記(4)で生成した2つのグラフ上で、それぞれ、例えば最小二乗法で近似直線を得て、プロットされた各対応関係と近似直線との乖離の程度を求めて、直線近似の度合いがより高い方のグラフを選択する。最小二乗法は、残差(実測の応答時間 - $f(x)$)の二乗和が最小となる $f(x)$ を求める計算方法である。

20

【0080】

図11は、直線近似および同時実行数の選択を説明するための図である。

【0081】

図11に示すグラフによって、左側のグラフの方が直線近似の度合いが高いと判断されたとする。この場合に、左側のグラフが対応する第1の組データの同時実行数の種別が選択され、応答時間モデルとして「全ジョブ合計」の同時実行数が選択される。

【0082】

ジョブが、例えばCPUのように全ジョブ種共通で使用する資源に相関する場合は、左側のグラフ、すなわち全ジョブ合計の平均同時実行数による対応関係の近似の度合いが高くなる。一方、ジョブが、例えば、トランザクションの排他ロックのようにジョブ種固有で使用する資源に相関する場合は、右側のグラフ、すなわちジョブ種単独の同時実行数による対応関係のほうが近似の度合いが高くなる。

30

【0083】

(6) 各ジョブ種について、選択したグラフにおいて、近似直線(平均値)からの応答時間の標準偏差を計算し、さらに、その信頼区間を計算し、信頼区間の上限および下限を計算して応答時間の正常範囲の上限/下限とする。

【0084】

図12は、グラフ上での信頼区間による応答時間の正常範囲の設定例を示す図である。

40

【0085】

選択した同時実行数によるグラフをもとに、標準偏差は、直線近似の際に求めた二乗和を、(標本数 - 1)で割って平方根をとったもので求め、信頼区間(99.99%)は、「近似直線(平均値) ± 信頼水準の定数(3.89) × 標準偏差」で求める。信頼度の定数(3.89)は、99.99%の信頼水準における値であり、標準正規分布表より求めることができる。

【0086】

応答時間の正常範囲の上限値は「 $f(x) + 3.89 \times \text{標準偏差}$ 」で求め、下限値は「 $f(x) - 3.89 \times \text{標準偏差}$ 」で求める。

50

【0087】

(7) 各ジョブ種の応答時間の正常範囲の上限/下限を示す応答時間モデルを応答時間モデル記憶部19に記憶する。

【0088】

図13は、応答時間モデルの例を示す図である。

【0089】

応答時間モデルとして、ジョブ種毎に、同時実行数の種類および応答時間の正常範囲を求める式を記録する。「同時実行数の種類」は、上記(5)で選択された同時実行数の種類(全ジョブ合計、ジョブ種単独)を示す。正常範囲を求める式 $f(x)$ の x は、選択した同時実行数の種類で定義された同時実行数である。ここでは、応答時間の正常範囲を求める式として、正常範囲の上限値を求める式が記録されている。

10

【0090】

次に、応答時間監視装置1の診断部15の診断処理(診断フェーズ)をより詳しく説明する。

【0091】

診断部15は、予め設定された一定時間毎に、分類済みメッセージ記憶部18に記憶された一定の時間に収集されたプロトコルメッセージのデータストリームを入力し、以下の処理を行う。

【0092】

(1) ジョブ毎に分類したプロトコルメッセージから紐付けたリクエストのプロトコルメッセージの受信時刻を開始時刻、レスポンスのプロトコルメッセージの受信時刻を終了時刻として各ジョブの応答時間を計算する。

20

【0093】

(2) 全ジョブ合計の平均同時実行数と、ジョブ種毎に、ジョブ種単独の平均同時実行数および各ジョブ種の平均応答時間とを計算する。

【0094】

(3) 各ジョブ種について、応答時間モデル記憶部19に記憶された当該ジョブ種の応答時間モデルを参照し、「同時実行数の種類」に設定された種別(全ジョブ合計またはジョブ種単独)の平均同時実行数で「応答時間の正常範囲を求める式」を用いて当該ジョブ種の応答時間の閾値(上限値、下限値)を求める。

30

【0095】

(4) 各ジョブ種について、計算した平均応答時間が、上記(3)で求めた閾値(上限値)を超えているかを判断する。

【0096】

(5) 計算した平均応答時間が閾値を超えている場合に、そのジョブ種を、正常ではない応答時間(遅すぎる応答時間)を記録したジョブ種として出力する。

【0097】

上記(1)~(5)の診断処理が一定時間毎に実行され、この間に遅い応答時間を記録したジョブ種のリストがアラートとして出力される。

【0098】

図14および図15は、一実施例における学習処理の処理フロー例を示す図である。

40

【0099】

学習部14は、分類済みメッセージ記憶部18から一定時間に収集されたデータを学習データとして入力し、入力データを細分化区間に分割する(ステップS10)。

【0100】

学習部14は、細分化区間が残っている限りループ処理を行う(ステップS11、S17)。

【0101】

学習部14は、その細分化区間における、全ジョブ合計の平均同時実行数を計算する(ステップS12)。

50

【0102】

さらに、学習部14は、ジョブ種毎にループ処理を行う(ステップS13、S16)。

【0103】

学習部14は、その細分化区間における、選択されたジョブ種単独の平均同時実行数を計算し(ステップS14)、さらに、選択されたジョブ種の平均応答時間を計算する(ステップS15)。

【0104】

さらに、学習部14は、ジョブ種毎にループ処理を行う(ステップS18、S24)。

【0105】

学習部14は、全ジョブ合計の同時実行数と選択されたジョブ種の平均応答時間との関係を直線近似し(ステップS19)、選択されたジョブ種単独の同時実行数とそのジョブ種の平均応答時間との関係を直線近似する(ステップS20)。

10

【0106】

そして、学習部14は、2つの直線近似の中で近似の度合いが高い方で用いた同時実行数の種別(全ジョブ合計、ジョブ種単独)を選択する(ステップS21)。

【0107】

さらに、学習部14は、各細分化区間について、近似直線(平均値)からの応答時間の差を求め、そこから標準偏差を計算し(ステップS22)、統計上の信頼基準を用いて応答時間の信頼区間を計算し、その上限/下限(直線)を求める式を応答時間の正常範囲(上限/下限)を求める式として応答時間モデル記憶部19に記録し(ステップS23)、

20

【0108】

図16は、一実施例における応答時間の診断処理の処理フロー例を示す図である。

【0109】

診断部15は、停止コマンドを受けるまで無限にループ処理を繰り返す(ステップS30、S33)。診断部15は、分類済みメッセージ記憶部18から一定時間分の入力データをデータストリームから読み込み(ステップS31)、個々の時間分の診断処理を行い(ステップS32)、停止コマンドにより処理を終了する。

【0110】

図17は、一実施例におけるステップS32の診断処理のより詳細な処理フロー例を示す図である。

30

【0111】

診断部15は、入力データの区間における、全ジョブ合計の平均同時実行数を計算し(ステップS320)、ジョブ種毎にループ処理を行う(ステップS321、S329)。

【0112】

診断部15は、入力データの区分における、選択されたジョブ種の平均応答時間を計算し(ステップS322)、選択されたジョブ種について、学習部14で定められた同時実行数の種類が「ジョブ種単独」であるかを判定する(ステップS323)。同時実行数算出時の種類が「ジョブ種単独」であれば(ステップS323のY)、入力データの区間における、選択されたジョブ種単独の平均同時実行数を計算し(ステップS324)、同時

40

【0113】

実行数の種類が「ジョブ種単独」でなければ(ステップS323のN)、入力データの区間における、全ジョブ合計の平均同時実行数を計算し、同時実行数として利用する(ステップS325)。

診断部15は、求めた同時実行数に応じた応答時間の閾値(上限値)を応答時間モデルから算出し(ステップS326)、選択されたジョブ種の平均応答時間が閾値(上限値)を超えているかを判定する(ステップS327)。選択されたジョブ種の応答時間が閾値を超えていれば(ステップS327のY)、選択されたジョブ種を応答時間が異常に長いジョブ種として出力する(ステップS328)、選択されたジョブ種の応答時間が閾値を超えていなければ(ステップS327のN)、全てのジョブ種についてのループ処理終了

50

後に処理を終了する。

【0114】

以上説明したように、応答時間監視装置1によれば、学習フェーズにおいて、ジョブ種毎に応答時間が通常取り得る値（正常範囲）を反映した応答時間モデルを得ることによって、診断フェーズにおいて、ジョブ種に応じた応答時間の正常範囲を設定するため、応答時間の自然なばらつきの範囲であればアラートを出力しない。

【0115】

したがって、応答時間監視装置1がアラートを出力した場合は、応答時間が同時実行数の増加による自然な増加の範囲ではなく、何らかの問題が発生していることがわかる。

【0116】

また、応答時間監視装置1によれば、ジョブ種毎の同時実行数と応答時間の関係を直線近似することにより学習するため、正常な状態（同時実行数が少ない）時の学習データだけを使って応答時間をモデル化することが可能である。

【0117】

本発明の一態様として開示した応答時間監視装置1によれば、次のような効果がある。

【0118】

・対処が不要な自然な応答時間の増加についてアラートしないため、無駄なアラートを減少させることができる。

【0119】

・アラートが本当に問題な状況によるものなのかを診断する手間を省くことができる。例えば、従来行っていたような、普段の応答時間と比較して異常であるかを調査したり、同時実行ジョブ数と突き合わせて、一時的なジョブの混雑なのかを調査したりする必要がなくなる。

【0120】

・監視対象で資源が飽和した状況のデータが不要であり、通常負荷の場合のデータで学習し応答時間モデルを生成することができる。

【0121】

これらの効果を奏することにより、応答時間監視装置1は、応答時間の監視をより精度良く行うことができ、コンピュータ上で提供されるサービスの性能安定化をはかることができる。

【0122】

開示する応答時間監視装置1を構成する要素は、任意の組合せで実現されてもよい。複数の構成要素が1つの部材として実現されてもよく、1つの構成要素が複数の部材から構成されてもよい。また、応答時間監視装置1は、上述した実施形態に限定されず、本発明の要旨を逸脱しない範囲において各種の改良および変更を行ってもよいことは当然である。

【符号の説明】

【0123】

- 1 応答時間監視装置
- 11 パケット収集部
- 12 プロトコル解析部
- 13 ジョブ種分類部
- 131 ジョブ種分類ルール
- 14 学習部
- 15 診断部
- 16 結果出力部
- 17 メッセージ記憶部
- 18 分類済みメッセージ記憶部
- 19 応答時間モデル記憶部
- 2 監視対象システム

10

20

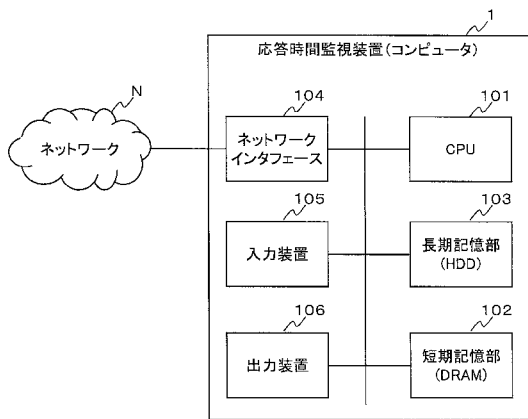
30

40

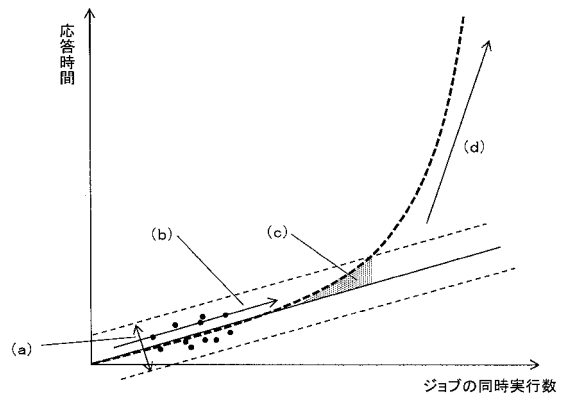
50

- 2 1 Webサーバ
- 2 2 DBサーバ
- 3 ネットワークスイッチ
- 4 プロキシ
- 5 インターネット
- 6 クライアント
- 7 性能監視サーバ

【 図 1 】



【 図 2 】

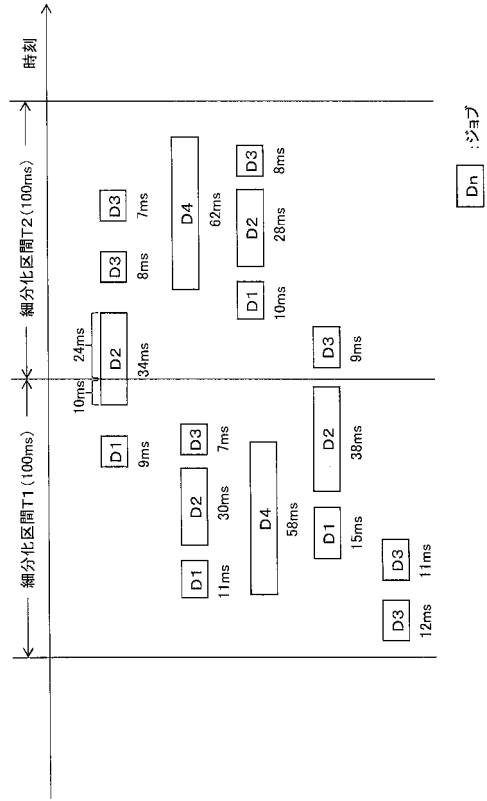


【 図 7 】

131

ジョブ種	コマンド名	ジョブ内容(SQL文を變換ルールで抽象化したもの)
DB-1	EXECREADREQ	SELECT .. FROM STORIES, USERS WHERE ..
DB-2	EXECREADREQ	SELECT .. FROM COMMENTS WHERE .. ORDER BY ..
DB-3	EXECREADREQ	SELECT .. FROM COMMENTS, USERS WHERE ..
DB-4	EXECREADREQ	SELECT .. FROM COMMENTS WHERE ..
DB-5	EXECREADREQ	SELECT .. FROM CATEGORIES
DB-6	EXECREADREQ	SELECT .. FROM STORIES WHERE .. ORDER BY ..
DB-7	EXECREADREQ	SELECT .. FROM OLD_COMMENTS WHERE ..
DB-8	EXECREADREQ	SELECT .. FROM OLD_COMMENTS WHERE .. ORDER BY ..
DB-9	EXECREADREQ	SELECT .. FROM OLD_COMMENTS, USERS WHERE ..
DB-10	EXECREADREQ	SELECT .. FROM OLD_STORIES, USERS WHERE ..

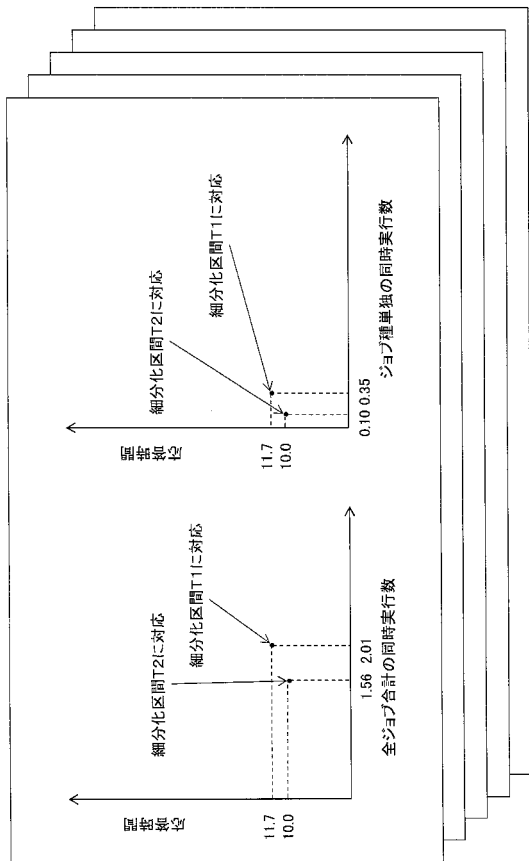
【 図 8 】



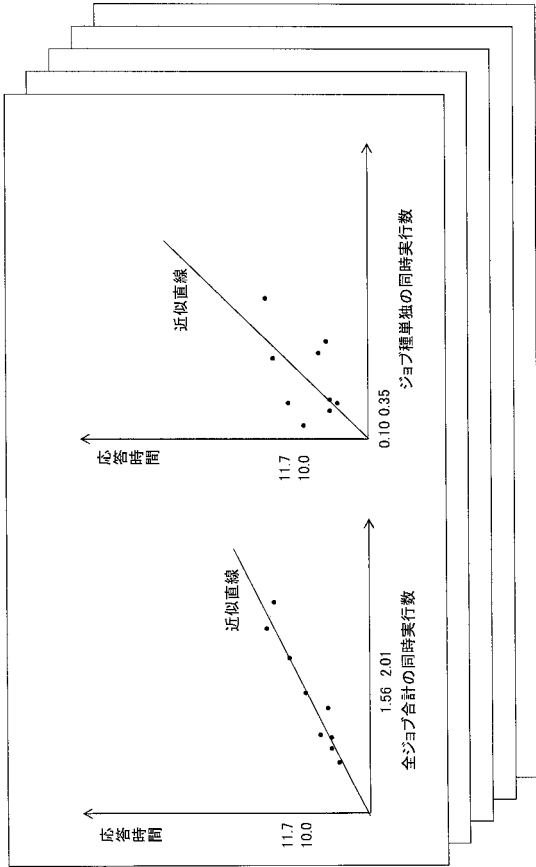
【 図 9 】

	全ジョブ合計 の平均同時 実行数	D1の 平均同時 実行数	D2の 平均同時 実行数	D3の 平均同時 実行数	D4の 平均同時 実行数	D1の 平均応答 時間(ms)	D2の 平均応答 時間(ms)	D3の 平均応答 時間(ms)	D4の 平均応答 時間(ms)
細分化区間 T1	2.01	0.35	0.78	0.30	0.88	11.7	34.0	10.0	58.0
細分化区間 T2	1.56	0.10	0.52	0.32	0.62	10.0	30.5	8.0	62.0

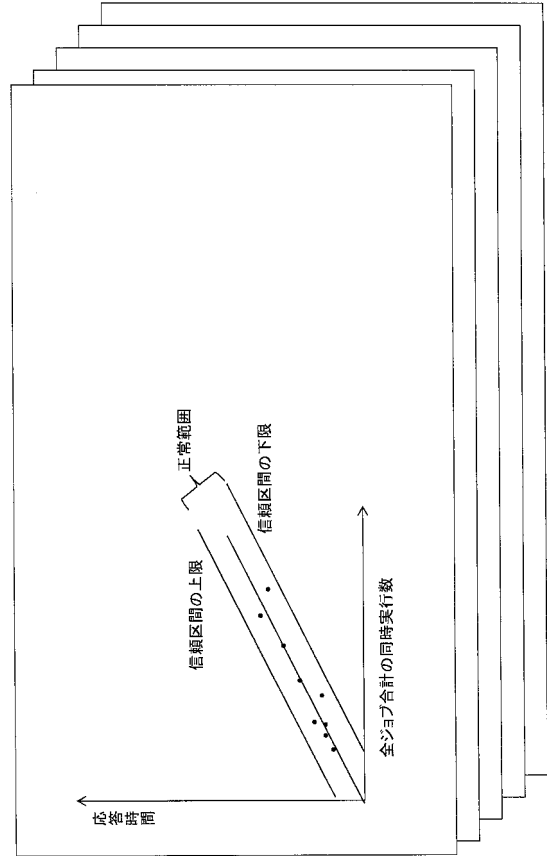
【 図 10 】



【図 1 1】



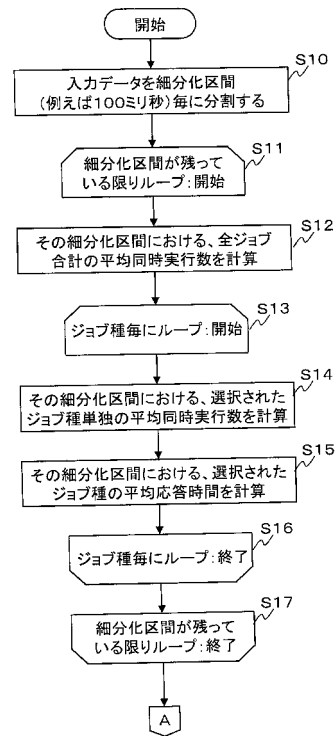
【図 1 2】



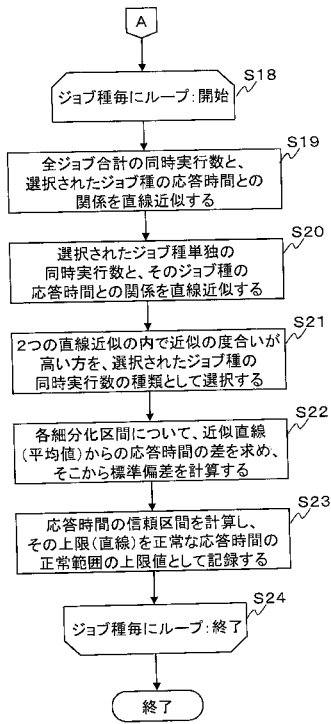
【図 1 3】

ジョブ種	同時実行数の種類	応答時間の正常範囲を求める式 (x : 左の列で定義された同時実行数)
DB-10	全ジョブ合計	$f(x) = 0.025x + 0.0125$
DB-460	ジョブ種単独	$f(x) = 52x + 0.016$

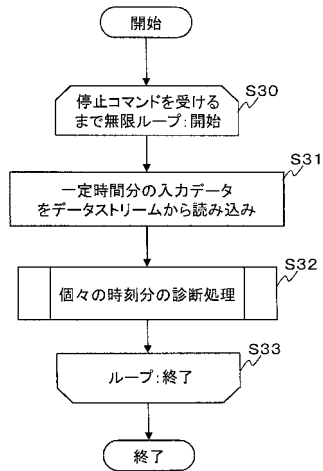
【図 1 4】



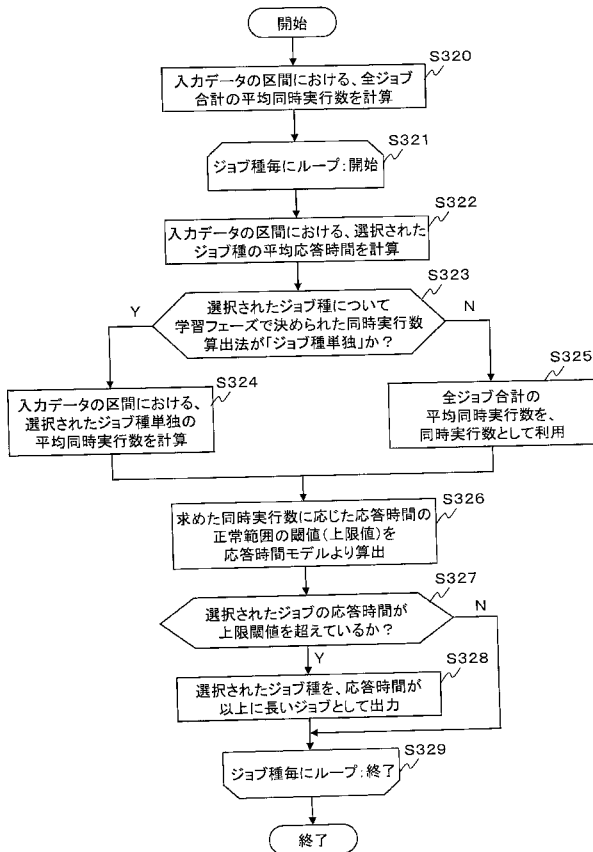
【 図 1 5 】



【 図 1 6 】



【 図 1 7 】



フロントページの続き

(72)発明者 松原 正純

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

Fターム(参考) 5B042 GA23 GB01 JJ08 JJ20 JJ29 MA14 MC21 MC33